Tidy data ( Dữ liệu gọn gàng):

+Biến dữ liệu đầu vào thành các chuẩn và thông tin không khác nhau chỉ biến đổi hình dạng cấu trúc cột và dòng.

+Để cho người dùng có thể dễ dàng phân tích sữ dụng.

Khái niệm: là chuyển dữ liệu thành các dạng khác nhau nhiều hàng hơn hoặc nhiều cột hơn. Tùy vào người sữ dụng.

Định nghĩa:

Mỗi biến là môt cột.

Mỗi hàng là 1 quan sát k giống nhau

Mỗi cột khong bị vướng với nhau

Mỗi ô là một giá trị đơn

Tidy data là đạt dạng chuẩn NF3

Trong tidy data trường cột biến đều giống nhau